

UNIVERSIDADE ESTADUAL DE CAMPINAS
SISTEMA DE BIBLIOTECAS DA UNICAMP
REPOSITÓRIO DA PRODUÇÃO CIENTÍFICA E INTELECTUAL DA UNICAMP

Versão do arquivo anexado / Version of attached file:

Versão do Editor / Published Version

Mais informações no site da editora / Further information on publisher's website:

https://ogst.ifpenergiesnouvelles.fr/articles/ogst/full_html/2019/01/ogst190090/ogst190090.html

DOI: 10.2516/ogst/2019045

Direitos autorais / Publisher's copyright statement:

©2019 by EDP Sciences. All rights reserved.

DIRETORIA DE TRATAMENTO DA INFORMAÇÃO

Cidade Universitária Zeferino Vaz Barão Geraldo

CEP 13083-970 – Campinas SP

Fone: (19) 3521-6493

<http://www.repositorio.unicamp.br>

Integration of geostatistical realizations in data assimilation and reduction of uncertainty process using genetic algorithm combined with multi-start simulated annealing

Célio Maschio* and Denis José Schiozer

Center for Petroleum Studies – CEPETRO, University of Campinas – Unicamp, Campinas, SP 13083-970, Brazil

Received: 27 March 2019 / Accepted: 29 July 2019

Abstract. This paper introduces a new methodology, combining a Genetic Algorithm (GA) with multi-start simulated annealing to integrate Geostatistical Realizations (GR) in data assimilation and uncertainty reduction process. The proposed approach, named Genetic Algorithm with Multi-Start Simulated Annealing (GAMSSA), comprises two parts. The first part consists of running a GA several times, starting with certain number of geostatistical realizations, and the second part consists of running the Multi-Start Simulated Annealing with Geostatistical Realizations (MSSAGR). After each execution of GA, the best individuals of each generation are selected and used as starting point to the MSSAGR. To preserve the diversity of the geostatistical realizations, a rule is imposed to guarantee that a given realization is not repeated among the selected individuals from the GA. This ensures that each Simulated Annealing (SA) process starts from a different GR. Each SA process is responsible for local improvement of the best individuals by performing local perturbation in other reservoir properties such as relative permeability, water-oil contact, etc. The proposed methodology was applied to a complex benchmark case (UNISIM-I-H) based on the *Namorado Field*, located in the Campos Basin, Brazil, with 500 geostatistical realizations and other 22 attributes comprising relative permeability, oil-water contact, and rock compressibility. Comparisons with a conventional GA algorithm are also shown. The proposed method was able to find multiple solutions while preserving the diversity of the geostatistical realizations and the variability of the other attributes. The matched models found by the GAMSSA method provided more reliable forecasts when compared with the matched models found by the GA.

Nomenclature

AQD	Acceptable Quadratic Distance	k_{rwi}	Maximum relative permeability for water at residual oil saturation
BHP	Bottom-Hole Pressure	k_{rowi}	Maximum relative permeability for oil at connate water saturation
cf	Crossover fraction	$k_{x,y,z}$	Absolute permeability in x, y and z directions
C_p	Constant used to compute AQD	LD	Linear Deviation
CPOR	Rock Compressibility	m	Maximum number of GA executions
G	Number of initial geostatistical realizations used in the proposed workflow	MSSAGR	Multi-Start Simulated Annealing with Geostatistical Realizations
GA	Genetic Algorithm	n_s	Number of simulated annealing instances
GAMSSA	Genetic Algorithm with Multi-Start Simulated Annealing	N_{obs}	Number of observed data for a given data series
GR	Geostatistical Realization	N_p	Cumulative oil production
g	Sub-set of geostatistical realizations selected after the execution of each GA	NQD	Normalized Quadratic Distance
Hist	Observed data	NQDS	Normalized Quadratic Distance with Sign
		pk_{rw}	Exponent of relative permeability for water
		pk_{row}	Exponent of relative permeability for oil
		P_m	Probability of mutation
		Q_o	Oil rate
		Q_w	Water rate

* Corresponding author: celio@cepetro.unicamp.br

R	Number of reservoir output
SA	Simulated Annealing
T	Temperature
Tol	Tolerance used to compute AQD
woc	Water-Oil Contact
x	A generic candidate from SA
y	A n -dimensional vector composed by -1 , 0 and 1

Greek letter

δ	Random number drawn from a uniform distribution $U[0,1]$
----------	--

1 Introduction

History Matching (HM), or data assimilation, consists of using dynamic observed data (such as well rates and pressure, for example) to enhance the estimation of reservoir properties (attributes), such as porosity and absolute permeability, relative permeability, among others. The final goal of the process is to improve the predictive capacity of reservoir models reducing the uncertainty in the production forecast. One of the challenges of the history matching process is to match dynamic data while maintaining geological consistency of the matched models. To do so, it is desirable to carry out the HM process, integrated with the geostatistical modeling.

1.1 Integration of geostatistical modeling and history matching

The integration of geostatistical modeling and history matching leads to a complex optimization problem because the relationship between the input and output variables can be highly nonlinear. There are two main ways of performing such integration. The first comprises the coupling of geostatistical software modeling into the HM process. The other way consists of generating several geostatistical realizations and integrating these realizations into the HM process.

In the first line, [Maschio *et al.* \(2015\)](#) presented a framework for geostatistics-based history matching using genetic algorithm with adaptive bounds. [Avansi *et al.* \(2016\)](#) applied virtual wells based on pilot points to ensure realistic geomodels, maintaining geological continuity while fitting the reservoir models output to the dynamic data. The presented methodology effectively preserves the consistency of geological models during the history matching process. [Oliveira *et al.* \(2017\)](#) proposed a methodology to integrate regional multi-property image perturbation methods with a multivariate sensitivity analysis. The authors only used petrophysical properties (facies, porosity, and absolute permeability) during the history matching.

In the second line, [Bennett and Graf \(2002\)](#) generated multiple geostatistical realizations and performed history matching of each reservoir model that resulted from these realizations (using deterministic parameters, such as relative permeability). For the history matching, a gradient-based

optimization method was used. Gradient-based methods rapidly converge, but are easily trapped in local minima. History matching is a non-linear problem characterized by many local minima. Therefore, premature convergence of gradient-based, or any other optimization method, is an undesirable characteristic in the context of HM due to the non-uniqueness nature of the problem.

[Maschio *et al.* \(2008\)](#) also used multiple geostatistical realizations into the history matching using genetic algorithm and direct search optimization methods. To apply the direct search method, they firstly sorted the geostatistical realization according to the objective function (quality match) values. The fundamental difference between the present work and the work presented by [Maschio *et al.* \(2008\)](#) is that here we propose a novel hybrid optimization scheme combining genetic algorithm in a different way (see details in [Sect. 2](#)) with multi-start simulated annealing. On the other hand, [Maschio *et al.* \(2008\)](#) tested a basic version of Genetic Algorithm (GA) and compared it with a direct search method. It is important to highlight that the comparisons were done separately, that is, no hybrid scheme was proposed in [Maschio *et al.* \(2008\)](#).

Another possible way of incorporating geostatistics-based parameterization in HM is by using the gradual deformation approach as suggested by [Hu *et al.* \(2001\)](#). There is a fundamental difference between gradual deformation approach and the method proposed in our paper. The basis of the Gradual Deformation Method (GDM) is the generation of new realizations constraining the stochastic model by observed data, by deforming (perturbing) an initial realization of a stochastic model. Therefore, GDM is focused on the parameterization part of the HM process. In our optimization process (for history matching), we use a fixed set of geostatistical realizations previously generated in a geological software modeling. During the history matching process, the geostatistical realizations are not changed. They are properly combined with other uncertainties (such as relative permeability, for example) using the proposed hybrid optimization scheme (genetic algorithm and simulated annealing). Therefore, our approach is focused on the optimization part of the HM process.

[Guérillot and Roggero \(1995\)](#) proposed a method named Scenarios Matching (SM) and, later on, [Roggero and Guérillot \(1996\)](#) presented the Production Scenarios Test Method (in the same line as SM) by which new geological models are generated constrained to observe and add data. The basic principle, according to authors, consists in adding new data at future times corresponding to various forecasts scenarios and matching observed and added data. This method, also according to the authors, allows transferring the uncertainty associated to the geological model into an uncertainty on production forecast. Using Bayesian inversion theory, [Gautier and Noetinger \(2004\)](#) proposed a method to estimate geostatistical parameters such as correlation length and permeability variance from well test data. As in GDM, these approaches generate new realizations constrained in some way to observed data, while in our approach the set of geostatistical realization do not change and, in this way, we ensure that the prior geological description of the model is preserved.

1.2 Genetic algorithm

Genetic Algorithm (GA), initially introduced by [Holland \(1975\)](#), is a global optimization method based on natural selection. GA repeatedly modifies a population of individuals to form the next generations. The main driving processes (or operators) of a genetic algorithm are selection, mutation, and crossover.

The selection operator is responsible for selecting the individuals, called parents, which contribute to the population at the next generation. Roulette wheel is a common selection approach where an individual of the current population has a probability of being selected which is proportional to its fitness value. Mutation is the mechanism by which a new individual is created by the introduction of a gene structure that is different when compared to other individuals in the population. By applying random changes to individual parents to form children, the mutation mechanism creates a new offspring from one individual by changing one or more of its genes. Crossover combines two parents to form children for the next generation. It is the operation of redistributing genetic characteristics between two (parents) individuals of a population. The goal of the crossover operator is to retain good features from the previous generation. It enables the algorithm to extract the best genes from different individuals and recombine them into potentially superior children.

The mutation and crossover are essential to the genetic algorithm. Proper setting of these two operators is important to balance global and local search. Mutation is responsible for introducing diversity in the generations. A higher amount of mutation increases the population's diversity. On the limit, a genetic algorithm without crossover (only mutation) is essentially a random sampling process.

[Romero *et al.* \(2000\)](#) used genetic algorithms and geostatistical modeling to generate multiple reservoir descriptions. The authors claimed that the algorithm is relatively easy to implement and easily parallelized. [Romero and Carter \(2001\)](#) demonstrated that genetic algorithms can be applied to realistic reservoir characterization problem, with more than 1900 variables. The authors mentioned that the algorithm returns a set of solution (final population) from which a representative group can be selected and used to assess prediction uncertainty. [Floris *et al.* \(2001\)](#) applied genetic algorithm in a comparative study. They claimed the ease of parallelization and the ability of coping with multiple optima as positive features of the genetic algorithm. [Schulze-Riegert *et al.* \(2002\)](#) investigated Evolutionary Algorithms (EA) applied to the history matching of complex reservoirs. Methods to improve the convergence of EA by introducing prior information were applied. [Schulze-Riegert and Haase \(2003\)](#) combined evolutionary algorithms and local (gradient-based) optimization methods.

[Sayyafzadeh *et al.* \(2012\)](#) compared a single objective genetic algorithm and a multi-objective genetic algorithm. By making use of Pareto optimization (multi-objective optimization), a set of solutions named the Pareto front, which consists of non-dominated solutions, is provided. According to the authors, decisions can be made with more confidence using the proposed approach.

[Abdollahzadeh *et al.* \(2012\)](#) proposed a hybrid method combining GA with Estimation of Distribution Algorithms (EDA). The authors also applied GA and EDA separately for comparison. They stated that, depending on the problem type, GA, EDA, and Hybrid GA-EDA can achieve good quality matches while performing a global search in the parameter space.

[Abdollahzadeh *et al.* \(2013\)](#) presented an adaptive scheme which intelligently adapts the parameters of an evolutionary algorithm in order to balance the diversity of the population. The control parameter adapted was the number of selected individuals to generated solutions in each generation (which controls the balance between exploration and exploitation). To test the method, the authors used some analytical functions with two variables and a simple history matching synthetic problem with three uncertain parameters. Although the work brings some interesting ideas based on adaptation mechanism to balance diversity and convergence throughout evolution, the authors did not show the robustness of the method in more challenging history matching problems.

[Xavier *et al.* \(2013\)](#) compared different combinations of set of parameters for genetic algorithm for the automatic history matching, applied to a simple 2D reservoir model with a five-spot scheme. The authors exploited parallelism via the MPI library and a master-slave decomposition strategy.

More recently, [Chakra and Saraf \(2016\)](#) proposed an Adaptive Genetic Algorithm (AGA), which is a relatively new optimization technique with adaptive genetic operators. The initial population, generated from geostatistical software, is used by the GA operators (selection, crossover, and mutation) to generate new reservoir realizations. The authors considered only permeability realizations in their study, which can be a limitation in practical cases.

[Sanghyun and Stephen \(2018\)](#) compared GA and Particle Swarm Optimization (PSO). They observed that GA was able to update a large quantity of control parameters. However, GA is a costly algorithm which requires more computing time. But, according to the authors, using the GA as a follow-up method helped to reach better results.

1.3 Simulated annealing

Simulated Annealing (SA) is a stochastic optimization method inspired by the process of physical annealing in solids. The original idea of SA as an optimization method was proposed by [Kirkpatrick *et al.* \(1983\)](#). Differently from GA, which works based on a population of individuals, SA (in its basic form) is a sequential optimization algorithm. The algorithm starts from an initial guess (a point selected from the search space) and generates perturbations (sequentially) around the current point.

The key feature of SA is its ability to escape from local minimum. The algorithm computes the difference (ΔOF) between the OF values corresponding to the current and the previous solutions and generates a random number (δ) from a uniform distribution between zero and one. If equation (1) is met, the current solution is accepted. Otherwise, the previous one is maintained as the current solution. The process continues proposing a new candidate

by perturbing the current solution. Note that improved solutions are always accepted while worse solutions (uphill moves) are accepted according to an acceptance probability. T is a control parameter which corresponds to temperature in the analogy with physical annealing. Equation (1) shows that small increases in ΔOF are more likely to be accepted than large increases and that, the higher the value of T , the higher the acceptance probability:

$$\delta \leq e^{(-\Delta OF/T)}. \quad (1)$$

The main set of control parameters of a typical simulated annealing algorithm are: (1) related-temperature parameters, which involve the definition of an initial temperature and a cooling schedule, comprising a temperature length (number of iterations at a given temperature) and a cooling ratio (rate at which the temperature is reduced); (2) the neighborhood characteristic; and (3) stopping criterion, common in any optimization method. Conventional SA algorithm normally starts with a relatively high value of T , to avoid being prematurely trapped in a local optimum. The main drawback of conventional SA is the high number of iterations necessary to guide the search to good solutions. This issue may make its application prohibitive in problems where the evaluation of the OF requires a long computational time, which is normally the case in history matching applications.

Some studies from the literature have addressed the application of SA in history matching. Using simulated annealing as an optimization method, [Sultan et al. \(1994\)](#) proposed an automatic history matching process to solve an inverse problem. Two major advantages of SA were highlighted by the authors: (1) it can deal with a large number of reservoir parameters; and (2) it enables to integrate different disciplines such as geology, petrophysics, and reservoir engineering. A general overview of the application of global optimization methods, including simulated annealing, in the context of dynamic reservoir inverse problems and reservoir description was presented by [Ouenes et al. \(1994\)](#). The authors highlighted the relevance of these methods to overcome the uniqueness issue that arises from the reservoir dynamic inverse problem.

[Chen et al. \(2012\)](#) applied a Very Fast Simulated Annealing (VFSA) to the history matching problem. A comprehensive comparison of several stochastic data-integration algorithms for the joint history matching of production and time-lapse seismic data was presented by [Long et al. \(2012\)](#). The authors concluded that VFSA is an effective single-model based search technique.

[Maschio and Schiozer \(2018\)](#) presented a novel hybrid scheme combining Iterative Discrete Latin Hypercube (IDLHC), proposed by [Maschio and Schiozer \(2016\)](#), with multi-start simulated annealing, introducing the method Iterative Discrete Latin Hypercube with Simulated Annealing (IDLHCSA). After running the IDLHC, good candidate solutions are selected based on Euclidian distance to ensure the variability and avoid the exploration of a specific region (local minimum) of the search space. The selected candidate solutions are used as initial points for multiple instances of the simulated annealing. It is important to highlight that, although we also use simulated annealing in the present

paper, the methodological procedure proposed here is completely different from the methodology proposed by [Maschio and Schiozer \(2018\)](#). The first fundamental difference is that here we use a set of geostatistical realizations (previously generated in a geostatistical software modeling) as part of the history matching parameters, that is, the geostatistical realizations are one of the “dimensions” of the n -dimensional search space (see details in [Sect. 3](#)). On the other hand, in [Maschio and Schiozer \(2018\)](#), the parametrization was based on zonation, *i.e.*, the reservoir model was divided into different regions and multipliers were applied for petrophysical properties in each region, using only one geostatistical realization as basis. The second main difference is that here we use genetic algorithm (see details in [Sect. 2](#)) in conjunction with SA, while in [Maschio and Schiozer \(2018\)](#), SA was used in conjunction with IDLHC. Other details and particularities can be found in [Maschio and Schiozer \(2018\)](#).

1.4 Integration of genetic algorithm and simulated annealing

The integration of GA and SA to compose a hybrid optimization scheme (GA-SA) is a relatively new approach in the literature. The approach consists of nesting the SA within GA such that SA improves individuals from GA. Example of such approach can be found in [Li and Wei \(2008\)](#), [Chen and Shahandashti \(2009\)](#), and [Junghans and Darde \(2015\)](#). Merging SA within GA is an interesting procedure. However, SA is essentially a serial method and running SA many times (serially) requires a long computational time and may be prohibitive for application in history matching. Thus, our purpose is to divide the process into two parts, running GA and SA separately, using the best solutions from several executions of GA (first part) as starting points for multiple instances of SA (second part).

1.5 Objective and contributions

The objective of this paper is to present a robust methodology to integrate geostatistical realizations in data assimilation and reduction of uncertainty process combining genetic algorithm with multi-start simulated annealing. The key idea is to provide a framework for HM that preserves the geological consistency, which is a challenge in any HM process.

The main contributions and novelties of this paper are: (1) to the best of our knowledge, this is the first study that proposes the use of GA integrated with multi-start SA to the HM problem; (2) it is the first application of a hybrid GA-SA approach to integrate geostatistical realizations in the HM process; (3) to overcome the drawback of GA-SA proposed until now in the literature, we propose executing the GA and SA in two parts, as explained in [Section 2](#), while retaining the strengths of both methods.

2 Methodology

The proposed approach combines the best feature of both methods: the ability of the GA to introduce diversity on

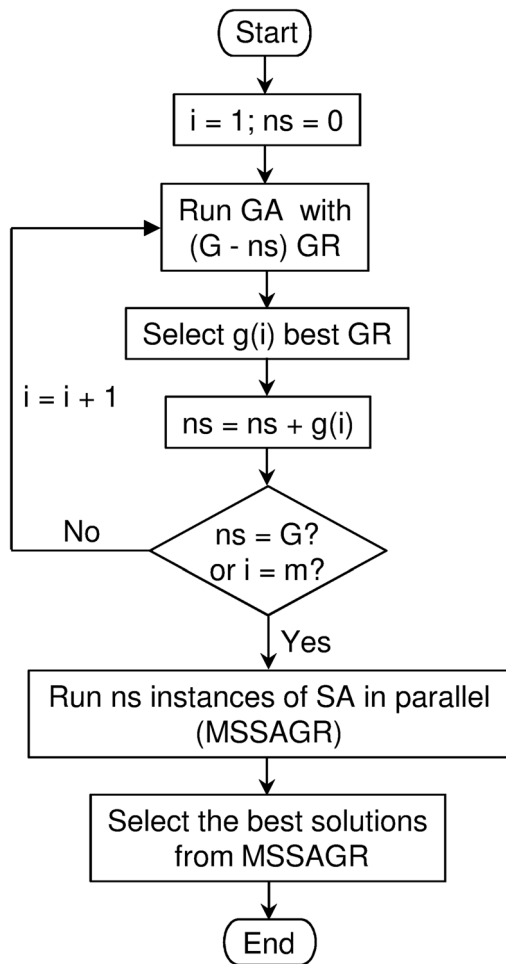


Fig. 1. Proposed workflow combining Genetic Algorithm (GA) with Multi-Start Simulated Annealing with Geostatistical Realization (MSSAGR).

the solutions and the abilities of SA to improve local solution and, at the same time, to escape from local minima. The workflow proposed in this work (Fig. 1) is divided into two parts. The first part consists of running the GA algorithm several times, starting with a number of geostatistical realizations (G) in the first execution. After each GA execution, a procedure (described in Sect. 2.1) is applied to select a quantity of individuals (g), which is a sub-set of G . The second part consists of running the Multi-Start Simulated Annealing with Geostatistical Realizations (MSSAGR). Each part of the process is explained in the following sections.

2.1 Optimization by GA and selection of the best individuals of each generation (first part)

In this work, an individual is composed by a Geostatistical Realization (GR), formed by spatial properties such as porosity and permeability, and other reservoir properties such as relative permeability, water-oil contact, etc. Each GR is identified by an integer number (index).

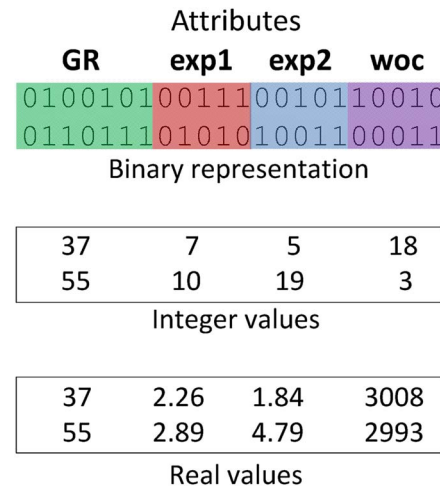


Fig. 2. Representation of individual in GA.

The search space for the GA is formed by the sequence of the geostatistical realizations and the other parameters. As the GR “dimension” is represented by integer indexes which do not represent any numeric sequence (or trend), gradient-based methods are not suitable for this kind of application. This is the motivation for using the GA method, which does not depend on descent direction.

The individuals of the GA are represented in this work by bit strings. Each attribute is represented by a number of bits which corresponds to an integer value (Fig. 2). The integer values corresponding to the real parameters (exp1, exp2 and woc in the schematic example in Fig. 2) are mapped to a vector of real values linearly distributed according to the range of the parameters (minimum a maximum values) and the number of intervals used to discretize the parameters. In each execution of GA, the number of generations and the number of individuals per generation are defined.

After each GA execution, the individuals of each generation are sorted according to the objective function values (Fig. 3) and a percentage of the best individuals are selected. From this percentage, repeated GR are filtered out in order to select individuals with different GR. In the schematic example of Figure 3, three individuals are selected from the first generation of GA1, two individuals are selected from the second generation and so on. Consider that the initial number of GR is 50 and that 15 GR were selected from the first execution of GA (GA1). Then, in the next execution of GA (GA2), the remaining 35 GR will be used in the process. This process continues until all GR have been selected or the maximum number of GA executions (m) has been reached. The selected individuals, which correspond to the number of SA instances (n_s), and its corresponding GR, are set as starting points for the MSSAGR.

To summarize, the objective of the GA in the proposed workflow is to perform a pre-optimization of each GR or, in other words, to find a good starting point for each GR to be used as input to the second part of the workflow (MSSAGR). The key role of the GA is to find good initial starting points (for each GR) for the MSSAGR.

Attributes					
GR	exp1	exp2	woc	OF	
GA1					
Generation 1 (GA1)					
1	36	2	5	15	1.70
2	36	3	8	17	2.50
3	36	3	8	17	2.60
4	48	10	16	3	2.80
5	48	12	17	4	2.95
6	12	8	11	16	4.15
...					
50	12	8	11	16	18.1
Generation 2 (GA1)					
1	27	4	9	14	1.54
2	27	7	6	17	2.15
3	36	3	8	17	2.50
4	35	17	8	4	2.70
5	78	12	17	4	2.85
6	12	4	5	16	6.15
...					
50	38	8	11	16	15.1
...					
Generation n (GA1)					
GA2					
...					
GAm					

Selected GR from each generation

Fig. 3. Schematic example of a sequence of GA executions with 50 individuals per generation.

2.2 Multi-start simulated annealing with geostatistical realizations (second part)

To explain how the MSSAGR takes place, consider the same example shown in Section 2.1. Each reservoir model is composed of a GR and the same three other attributes (exp1, exp2 and woc). Each SA instance is composed of a different GR and a given starting point (x_0) that comes from the selected individuals from the GA. The GR is fixed in each SA instance and the search space is formed by the other attributes (Fig. 4). In the example, 50 instances of the SA are run in parallel. For each instance, a given number of iterations are defined.

To perform the move mechanism of each SA instance, firstly a vector $\mathbf{s} = \{0, -1, 1\}$ is defined. Then, a vector \mathbf{y} of size n , where n is the number of dimensions, is sampled uniformly at random (with replacement) from the values in the vector \mathbf{s} . Thus, a new candidate (\mathbf{x}^{t+1}) is generated as follows:

$$\mathbf{x}^{t+1} = \mathbf{x}^t + \mathbf{y}, \quad (2)$$

where \mathbf{x}^t is a vector that represents the current state.

The vector \mathbf{s} determines the distance between the current state and the new one. As in this paper we use SA

MSSAGR			
Starting points			
GR from GA 1	exp1	exp2	woc
SA1: GR36	$x_{01} = [2$	5	15]
SA2: GR48	$x_{02} = [10$	16	3]
SA3: GR12	$x_{03} = [8$	11	16]
SA4: GR27	$x_{04} = [4$	9	14]
SA5: GR35	$x_{05} = [17$	8	4]
GR from GA 2			
SA6: GR10	$x_{06} = [7$	2	18]
SA7: GR42	$x_{07} = [16$	21	6]
SA8: GR39	$x_{08} = [9$	11	17]
...			
SA50: GR18	$x_{050} = [4$	18	14]

Fig. 4. Schematic representation of the Multi-Start Simulated Annealing with Geostatistical Realizations (MSSAGR).

as local search, large jumps are avoided by limiting the distance of the new candidate. The maximum allowed distance is the longest diagonal of a unit hypercube (note that the search space is formed by nodes in n -dimensional integer variables). As in GA (Sect. 2.1), the integer values generated by SA are mapped back into real values. Supposing that the current state in a two-dimensional case is represented by the vector $[5, 5]$, the possible new positions that a candidate can occupy are $\{[4, 4]; [4, 5]; [4, 6]; [5, 4]; [5, 6]; [6, 4]; [6, 5]; [6, 6]\}$.

2.3 Objective function

In this paper, the objective function is based on a Normalized Quadratic Distance (NDQ) computed according to equation (3) and used by several authors (Avansi *et al.*, 2016; Davolio and Schiozer, 2018; Maschio and Schiozer, 2016):

$$NQD = \frac{QD}{AQD}, \quad (3)$$

where,

$$QD = \sum_{i=1}^{N_{obs}} (\text{Sim}_i - \text{Hist}_i)^2, \quad (4)$$

and,

$$AQD = \sum_{i=1}^{N_{obs}} (\text{Tol} \times \text{Hist}_i + C_p)^2, \quad (5)$$

where Tol is a tolerance given by a percentage of the observed data (Hist) and C_p is a constant used to avoid excessive weight when the observed data is close to zero or to prevent division by zero when the observed data is equal to zero, which can occur with produced water rate. Sim represents the simulated results and N_{obs} is the

number of observed data for a given data series (water rate in a producer well, for example).

The global Objective Function (OF) used in the optimization processes is defined according to equation (6):

$$OF = \frac{1}{R} \sum_{j=1}^R NQD_j, \quad (6)$$

where R is the number of reservoir output.

In order to check if the variability of the solutions is biased or not, that is, if the ensemble of final solutions encompasses or not the observed data, it is interesting to compute the sign of the NQD, defining a Normalized Quadratic Distance with Sign (NQDS) as follows (Eq. (7)):

$$NQDS = \frac{LD}{|LD|} NQD, \quad (7)$$

where,

$$LD = \sum_{i=1}^{N_{obs}} (Sim_i - Hist_i). \quad (8)$$

Therefore, NQDS plot is an additional feature to assess the final results and to facilitate the visualization of the results.

3 Application

3.1 Case description

The proposed workflow was applied to the UNISIM-I-H benchmark case (UNISIM, 2015) which was built for history matching and uncertainty reduction purposes. Before the creation of UNISIM-I-H (Fig. 5), a very fine model (UNISIM-I-R) was created using real data from *Namorado Field* (Campos Basin, Brazil). Cores and well logging data from a total of 56 wells, 2D and 3D seismic data provided by the *Brazilian National Petroleum Agency – ANP* and also from *Petrobras* (released public data) were used to build the UNISIM-I-R. This high resolution geological model (3.5 million active cells) plays the role of our real reservoir.

Following a workflow normally applied in a real field development, firstly a geological model was built using the data of only four exploratory wells. This model was upscaled and a simulation model was created (UNISIM-I-D). This model was used to optimize a production strategy and, as a result, 14 producer (including the four defined for UNISIM-I-D) and 11 water injector wells were defined. Then, well log data from these 25 wells were extracted from the UNISIM-I-R and used to generate another geological model. Note that there is no relation from these 25 wells with the 56 wells used to build the UNISIM-I-R. After an upscaling process of this geological model, the UNISIM-I-H was built. More details can be found in Avansi and Schiozer (2015) and Avansi et al. (2016).

The data set of UNISIM-I-H, which is discretized into a corner point grid with $81 \times 58 \times 20$ cells with a total of

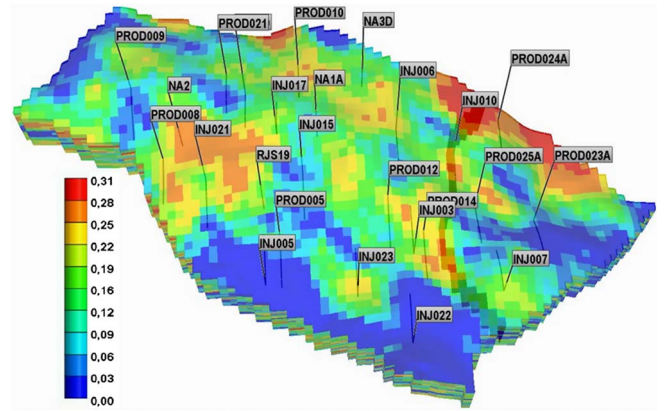


Fig. 5. UNISIM-I-H model: 3D porosity distribution and well locations.

36,739 active cells, is composed of 500 geostatistical realizations (e.g., Fig. 6) generated using a geostatistical modeling software. Each realization is composed of six spatial properties: (1) porosity, (2) horizontal permeability in x direction, (3) horizontal permeability in y direction, (4) vertical permeability (z direction), (5) net-to-gross ratio and (6) facies indicator. The six spatial properties are correlated, and for this reason, they are defined together as a single geostatistical realization.

During the historical period, the 14 producer wells are controlled by liquid rate and the 11 injector wells are controlled by water rate. The objective function is composed by a total of 78 components ($R = 78$ in Eq. (6)): 56 related to bottom-hole pressure, liquid, oil and water rate for the 14 producer wells and 22 related to bottom-hole pressure and water rate for the 11 injector wells. The tolerances (Tol and C_p) used to compute Acceptable Quadratic Distance (AQD, Eq. (5)) are shown in Table 1.

3.2 Pre-selection of geostatistical realizations

In this work, before the application of the proposed workflow, a sub-set of 100 geostatistical realizations (G) was selected. A simple way of doing this selection would be to run 500 models, each model being composed of one GR and the mean value of the other attributes. However, we performed a more elaborate procedure: instead of using just one level of the other attributes, we divided the range of each attribute into five levels and applied a Latin Hypercube Sampling method to combine each GR with the five levels of each attribute, totaling 2500 models. The motivation for this procedure is the fact that a given GR may have different HM quality when combined with different level of the other attributes. Thus, the proposed procedure is more robust in this sense.

3.3 Uncertain attributes

The uncertain attributes defined for the UNISIM-I-H for this study are presented in Table 2. There is a total of 23 attributes: 100 geostatistical realizations (selected

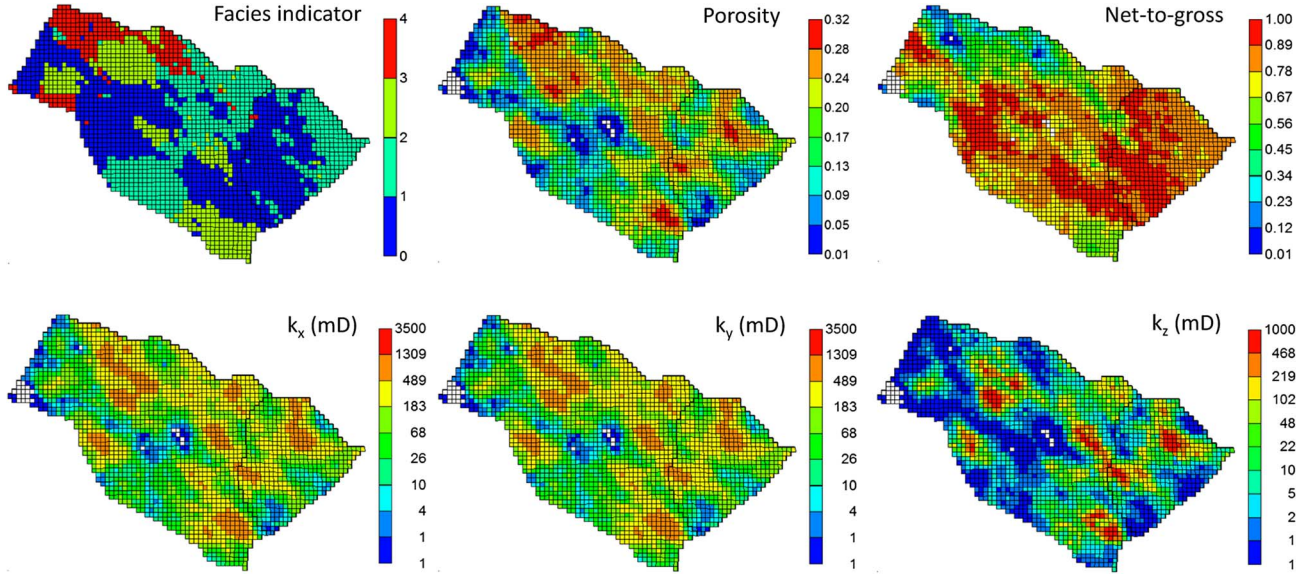


Fig. 6. Example of geostatistical realization (Layer 8) with: facies indicator, porosity, net-to-gross, horizontal permeabilities (k_x and k_y) and vertical permeability (k_z).

Table 1. Tolerance used to compute AQD (Eq. (5)).

Functions	Tol	C_p
Q_l	0.05	0
Q_o	0.1	0
Q_w	0.1	20 m ³ /d
Q_{wi}	0.05	0
BHP	0.1	0

AQD, Acceptable Quadratic Distance; BHP, Bottom-Hole Pressure; Tol, Tolerance.

according to Sect. 3.2), water-oil contact (block east), rock compressibility and 20 parameters related to the rock-fluid properties. For each of the four facies, we defined five parameters: exponent of relative permeability for water ($p_{k_{rw}}$) and oil ($p_{k_{row}}$), maximum relative permeability for water at residual oil saturation (k_{rwi}), maximum relative permeability for oil at connate water saturation (k_{rowi}) and exponent of capillary pressure (p_{pc}). These five parameters follow the power law described in CMG (2012). Except for the geostatistical realizations, all the other attributes were discretized into 31 levels. Maschio and Schiozer (2016) obtained good results for the same case using this number of levels.

3.4 Sensitivity analysis

To show the impact of the geostatistical realizations and the other attributes in the reservoir outputs, a sensitivity analysis was performed. To perform this analysis, four sets of models were generated and analyzed, as described below. For sets 1, 3 and 4, each attribute was discretized into five intervals.

1. In the first set (“GR atr”), 500 simulation models were generated combining the 100 GR with the other 22 attributes.
2. In the second (“GR”), the other attributes were fixed in the mean value, between the minimum and maximum value, and 100 simulation models were generated, each model corresponding to a GR.
3. In the third set (“atr1”), one GR was chosen among the one hundred, and 100 simulation models were generated, combining the 22 attributes using the DLHC method (Maschio and Schiozer, 2016).
4. The fourth (“atr2”) is similar to set 3, but another GR among the one hundred was chosen.

Figure 7 shows the NQDS plot for Q_o , Q_w and Bottom-Hole Pressure (BHP) for the producer and injector wells for the four sets of models. The first observation is that there is a similarity between the variability caused by the geostatistical realizations and the variability caused by the other parameters. Note that there is a similar spread between “GR”, “atr1” and “atr2” for most functions. This means that the influence of the other attributes and the GR are relatively well-balanced.

It is also possible to note that the variability caused by the combination of the GR with other attributes (“GR atr”) is similar to the variability caused by these attributes separately (“GR”, “atr1” and “atr2”) for many functions.

The third observation is that, mainly for Q_w and Q_o , the NQDS is relatively well-distributed, with negative and positive values, meaning that the model’s behavior is not biased with respect to the observed data.

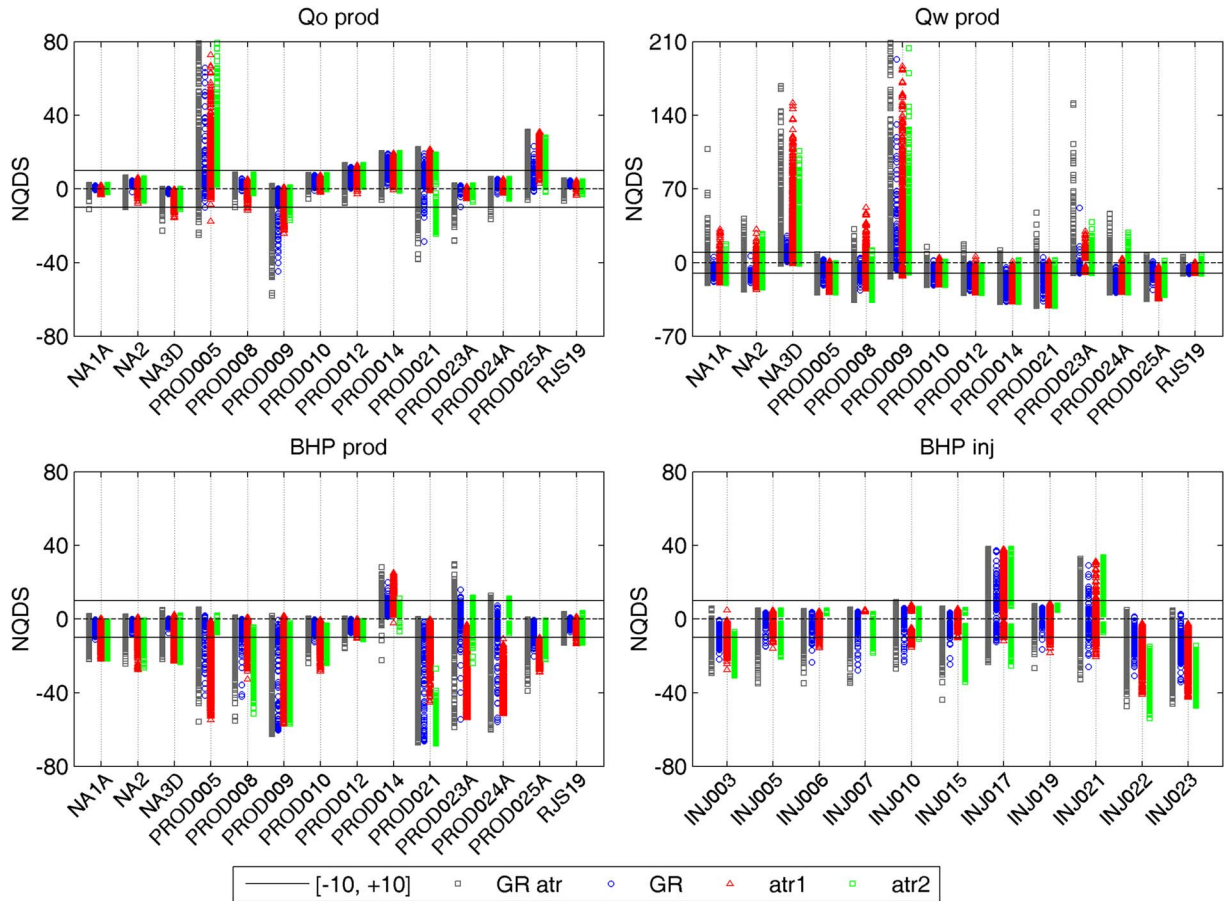
3.5 Control parameter setting for GA and SA

As the main control parameters for both GA and SA are problem-dependent, some preliminary tests are necessary

Table 2. Uncertain attributes defined for the UNISIM-I-H for this study.

Attribute	Description	Minimum	Maximum	Number of levels
GR	Six spatial properties (porosity, permeability in x , y and z directions, net-to-gross and facies indicator) for each GR	*	*	100
pk_{rw1} , pk_{rw2} , pk_{rw3} , pk_{rw4}	Exponent for water relative permeability (Facies 1–4)	1	6	31
pk_{row1} , pk_{row2} , pk_{row3} , pk_{row4}	Exponent for oil relative permeability (Facies 1–4)	1	6	31
k_{rwi1} , k_{rwi2} , k_{rwi3} , k_{rwi4}	Maximum oil relative permeability at residual oil saturation (Facies 1–4)	0.1	0.7	31
k_{rowi1} , k_{rowi2} , k_{rowi3} , k_{rowi4}	Maximum oil relative permeability at connate water saturation (Facies 1–4)	0.45	0.8	31
ppc1, ppc2, ppc3, ppc4	Exponent for capillary pressure (Facies 1–4)	3	11	31
woc	Water-Oil Contact (m)	3169	3179	31
CPOR	Rock Compressibility (kgf/cm^2) ⁻¹	10×10^{-6}	96×10^{-6}	31

* Each realization has a specific minimum and maximum value for each spatial property.
GR, Geostatistical Realizations.

**Fig. 7.** Impact of the geostatistical realizations and the other parameters in NQDS.

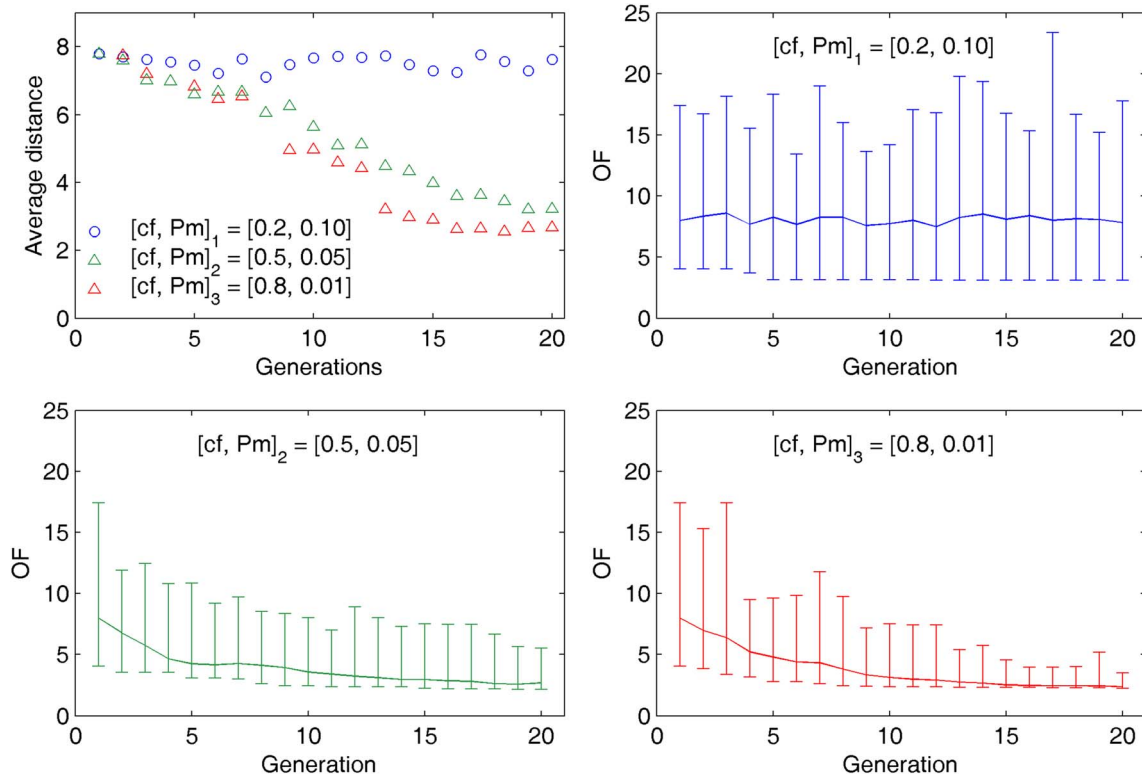


Fig. 8. Influence of the GA tuning parameters for the case studied.

for a proper choice. For the GA, three options varying the probability of mutation (P_m) and crossover fraction (cf) were tested. According to [Chakra and Saraf \(2016\)](#), the optimal value for cf reported in the literature ranges between 0.5 and 1.0. Usually, the P_m ranges between 0.001 and 0.05. Three executions with 20 generations and 100 individuals per generation were carried out according to the following values: $[cf, P_m]_1 = [0.2, 0.10]$, $[cf, P_m]_2 = [0.5, 0.05]$ and $[cf, P_m]_3 = [0.8, 0.01]$, where the subscript number after the bracket indicates the execution. From execution 1 through 3, the amount of mutation decreases and, consequently, the amount of crossover increases.

The top-left plot in [Figure 8](#) shows the average distance between the individuals for the 20 generations corresponding to the three runs. Note that the case with the largest amount of mutation presents the highest average distance. As the crossover fraction is too low and the probability of mutation is too high, the diversity of the individual remains very high along all generations. This implies in a poor convergence of the algorithm as can be seen in the top-right plot, which shows the best, worst, and mean OF values for each generation. The average distance (diversity of the individuals) decreases as the crossover increases and the probability of mutation decreases. It is also possible to see an improvement in the convergence for runs two and three.

Based on the previous analysis, we chose $[cf, P_m] = [0.5, 0.05]$ for the probability of mutation (P_m) and crossover fraction (cf) for all GA runs. In the first part of the workflow, five executions of GA were carried out.

The number of generations was set to 20 and the number of individuals per generation was set to 100 for each execution.

The influence of initial temperature on SA behavior for the case studied is shown in [Figure 9](#). The lower the temperature, the closer the models will be. As the purpose of this paper consists of starting each SA process from a pre-optimized model provided by GA, we chose a moderate small value ($T_i = 1.0$), allowing each SA process to explore the regions in the vicinity of each starting point but, at the same time, avoiding that the method moves away from these regions quickly. In the second part of the proposed workflow, 100 instances of the SA were launched with 500 iterations for each one.

To compare with the proposed method (Genetic Algorithm with Multi-Start Simulated Annealing [GAMSSA]), we carried out an additional run of GA (called in this paper as conventional GA) with 300 generations and 200 individuals per generation. We chose these numbers (300×200) to result in the same number of simulations used in the proposed method. The same values of cf and P_m used in the first part of the workflow were used in this run.

4 Results

This section is organized as follows: firstly, we analyze the first part of the workflow which consists of running the GA several times ([Sect. 4.1](#)). In sequence, the second part of the methodology (MSSAGR) is analyzed ([Sect. 4.2](#)).

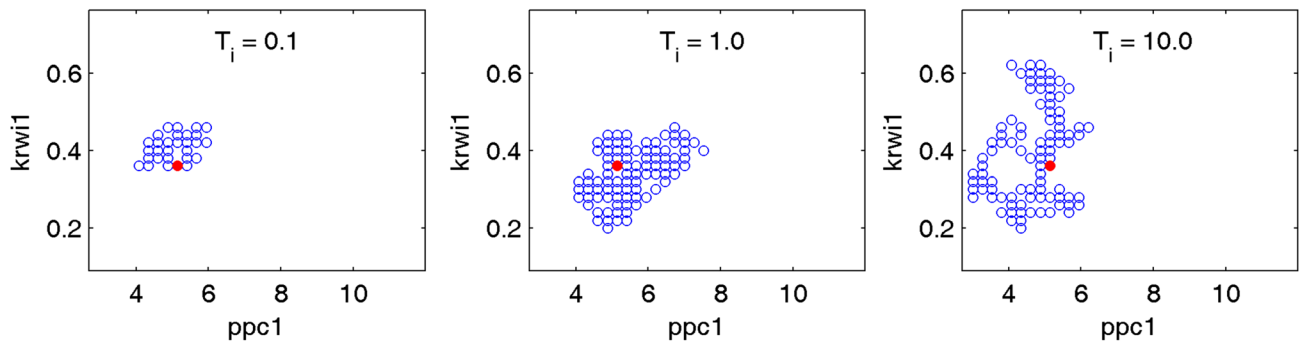


Fig. 9. Cross plot of a pair of attributes showing the influence of initial temperature on SA behavior for the case studied. The red color denotes the starting point and the assessed points are in blue.

Table 3. Number of geostatistical realizations selected in each of the five executions of GA.

GA execution	Number of selected GR	Number of GR used in the GA optimization
1	21	100
2	23	79
3	25	56
4	21	31
5	10	10

GA, Genetic Algorithm; GR, Geostatistical Realizations.

The final three subsections are dedicated to evaluate the history matching quality of the selected models (Sect. 4.3), the variability of the solutions (Sect. 4.4) and the analysis of the production forecast (Sect. 4.5).

4.1. Genetic algorithm (first part)

Table 3 shows the number of geostatistical realizations selected (according to the criterion described in the methodology) in each of the five executions of GA. Figure 10 shows the evolution of the objective function for two executions (GA1 and GA5). These plots show, in general, good convergence and an adequate variability along the generations.

4.2 Multi-start simulated annealing (second part)

The evolution of the SA algorithm can be seen in Figure 11 for 3 out of the 100 SA processes. Note that there is good convergence, showing significant improvement in the OF values. Figure 12 shows initial and lowest OF for the 100 SA instances. Overall, the initial OF values decreased roughly 50%. Considering the fact that each SA instance starts from a pre-optimized solution, this represents a good improvement.

4.3 Model selection and history matching quality

To select the final set of models for both methods, two steps were carried out: 1) for the conventional GA, the 20 best

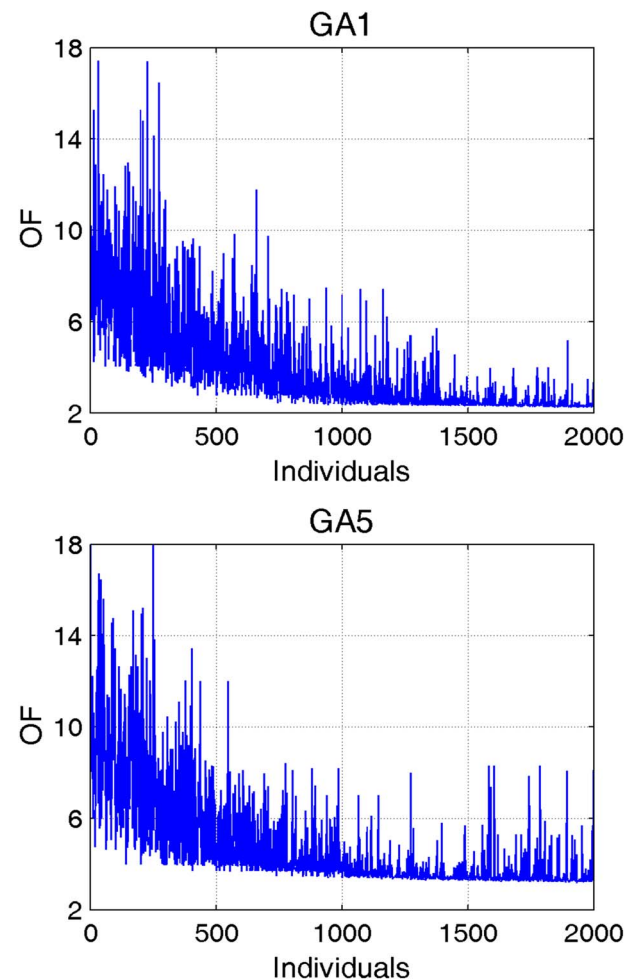


Fig. 10. Evolution of the objective function for two executions of GA (GA1 and GA5).

individuals (individuals with lower objective function) of each generation ($20 \times 300 = 6000$) were selected. For the GAMSSA, the 60 best solutions of each SA ($60 \times 100 = 6000$) were selected; 2) from the set of models selected in step 1, a filter of $|NQDS| = 10$ (excluding three problematic functions from the 78 defined for the case) was applied to

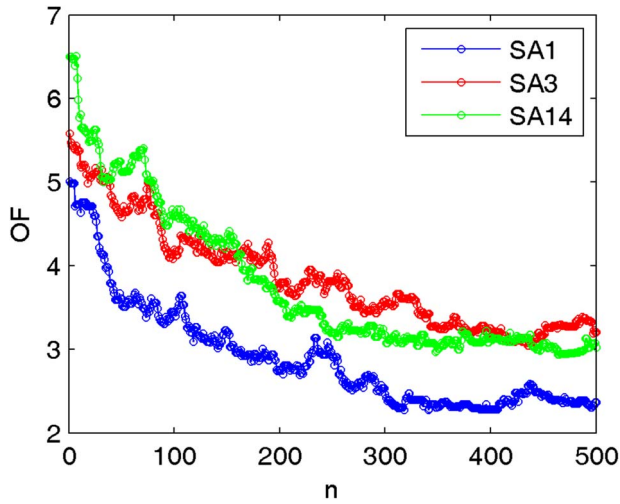


Fig. 11. Evolution of the OF for three instances of the SA algorithm.

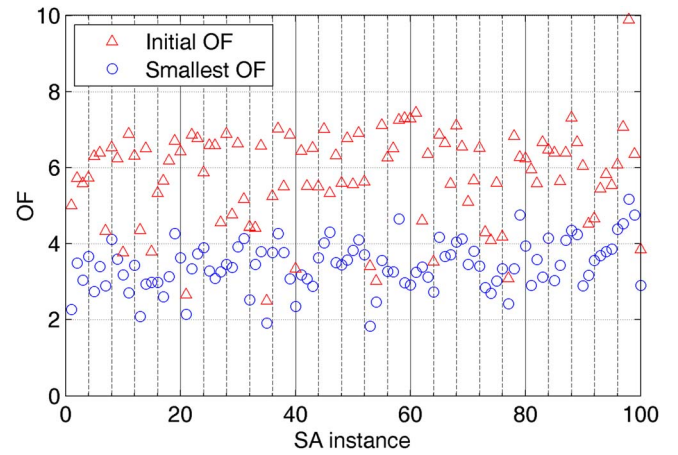


Fig. 12. Initial and smallest OF for the 100 SA instances.

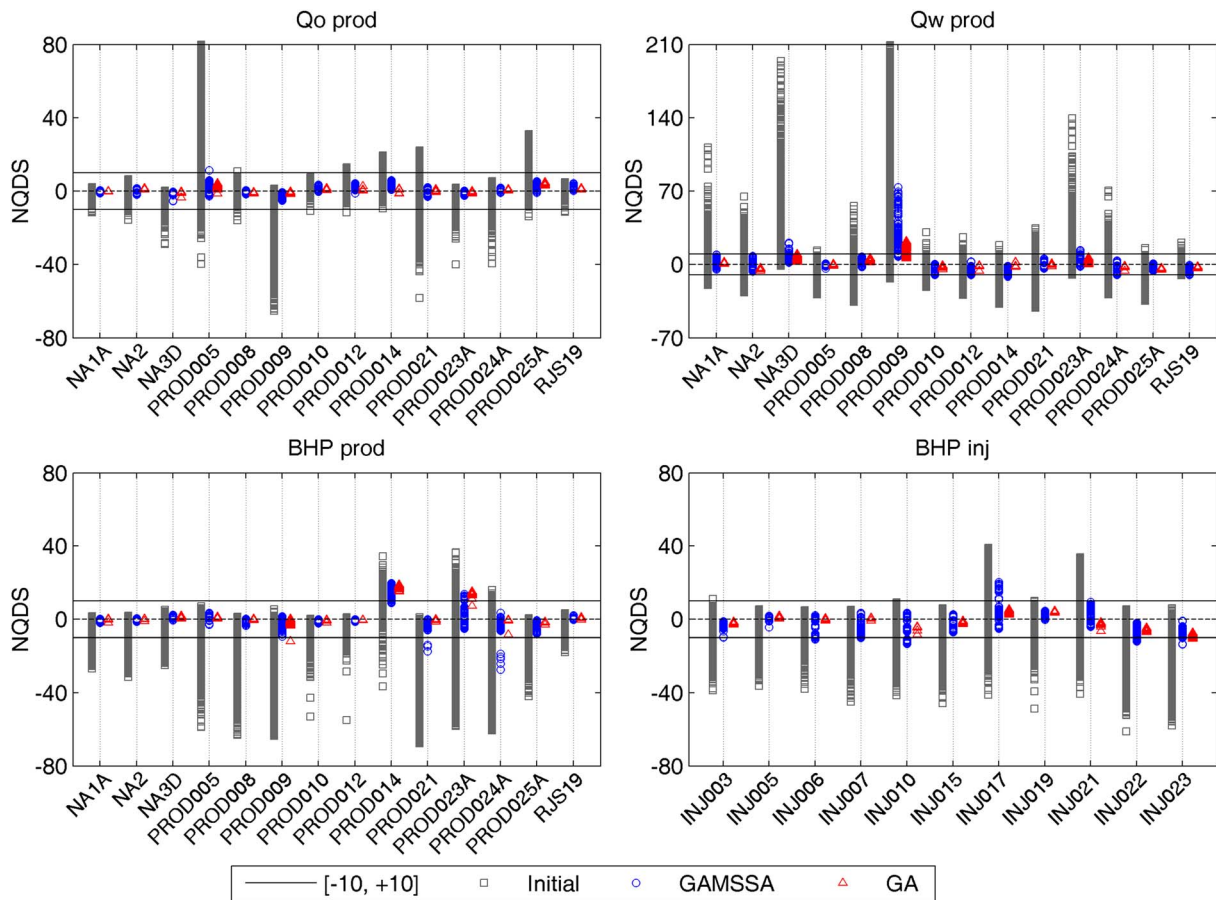


Fig. 13. Comparative plots of NQDS for the initial and the selected models from GAMSSA and GA.

select the final set of models for each method: 434 for GA and 194 for GAMSSA. The filter applied means that for the 434 models selected from the GA, and for the 194 models selected from the GAMSSA, at least 75 functions have, simultaneously, $|NQDS| \leq 10$.

Figure 13 shows comparative plots of NQDS for Q_o , Q_w , BHP for the producer wells and BHP for the injector wells. The gray square markers indicate the initial variability. An overall improvement in the quality of data match can be noted for both methods for most functions. Figure 14 shows

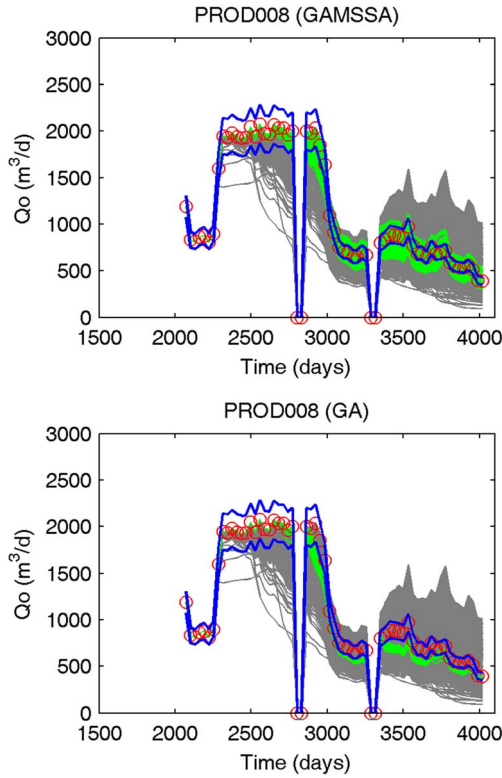


Fig. 14. Oil rate for the well PROD008 comparing GAMSSA and GA. Gray lines indicate the initial model dispersion, green lines are the selected matched models, red points are the observed data and the blue lines represent the tolerance.

the oil rate for one of the wells (PROD008) comparing GAMSSA and GA. Note that the selected models (green curves) from GAMSSA more properly encompass the observed data. It is important to analyze not only the quantity of the models, in terms of data match, but also the variability of the solutions, as described in [Section 4.4](#).

4.4 Variability of the solutions

[Figures 15 and 16](#) show cross plots for GAMSSA and GA, respectively. The gray points indicate the models selected according to the procedure described in [Section 4.3](#) and the blue (GAMSSA) and red (GA) points indicate the selected models using the filter based on $|NQDS| \leq 10$. The first observation is that the diversity of the solutions selected from GAMSSA is much larger than the solutions selected from GA. Analyzing the three plots on the top of [Figure 16](#), which shows the cross-plot between the GR and the exponent of Water Relative Permeability (pk_{rw}) curve for the three most important facies, we can note that GA concentrated the best solutions around just one GR (GR 35), while the best solutions from GAMSSA contemplate a larger number of GR. As further discussed in [Section 4.5](#), this different degree of diversity has a strong impact on the variability in the production forecast.

From the cross-plot between GR and pk_{rw} related to facies 2 and 3, it is possible to note that each GR, corresponding to the best solution for GAMSSA, exhibit different values of pk_{rw} . For example, looking at the top-right plot (facies 3), it can be noted that the best values of pk_{rw3} for the GR 33 are between 1.0 and 2.2, while for the GR 2, the best

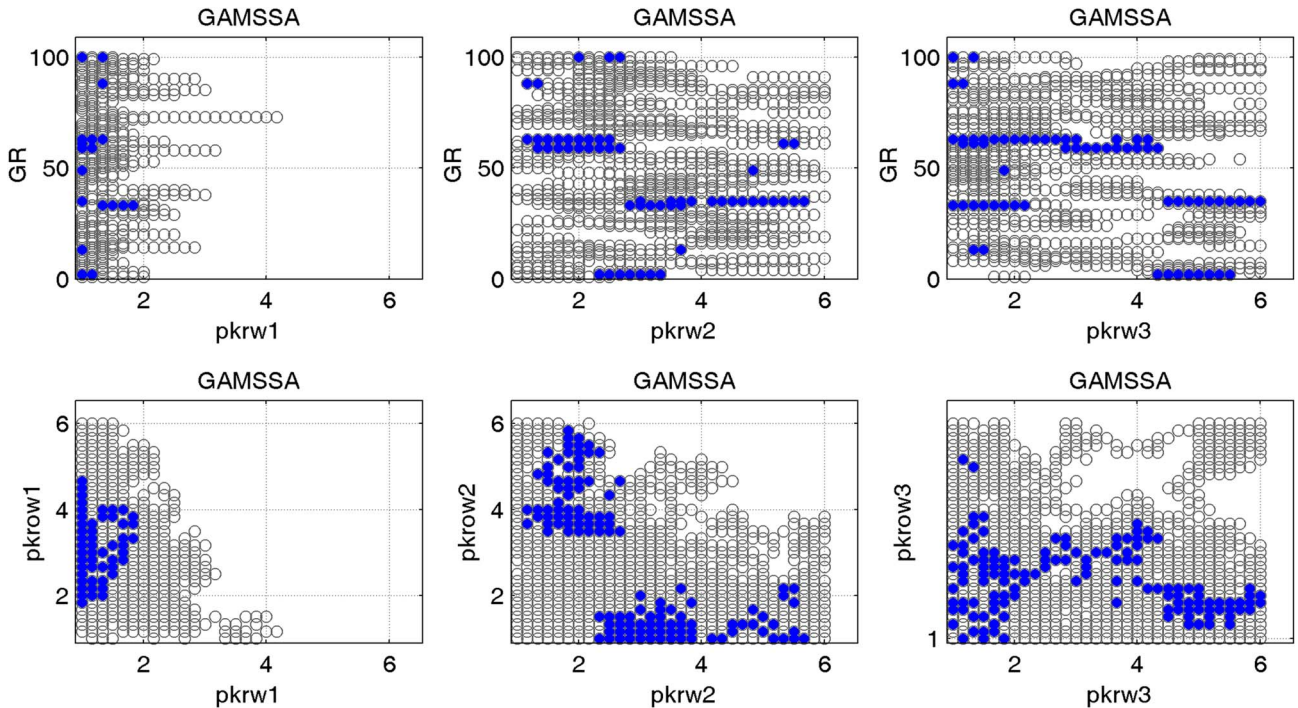


Fig. 15. Cross plots between GR and pk_{rw} (GAMSSA). Blue points denote models with $|NQDS| \leq 10$ and gray points indicate the models selected according to the procedure described in [Section 4.3](#).

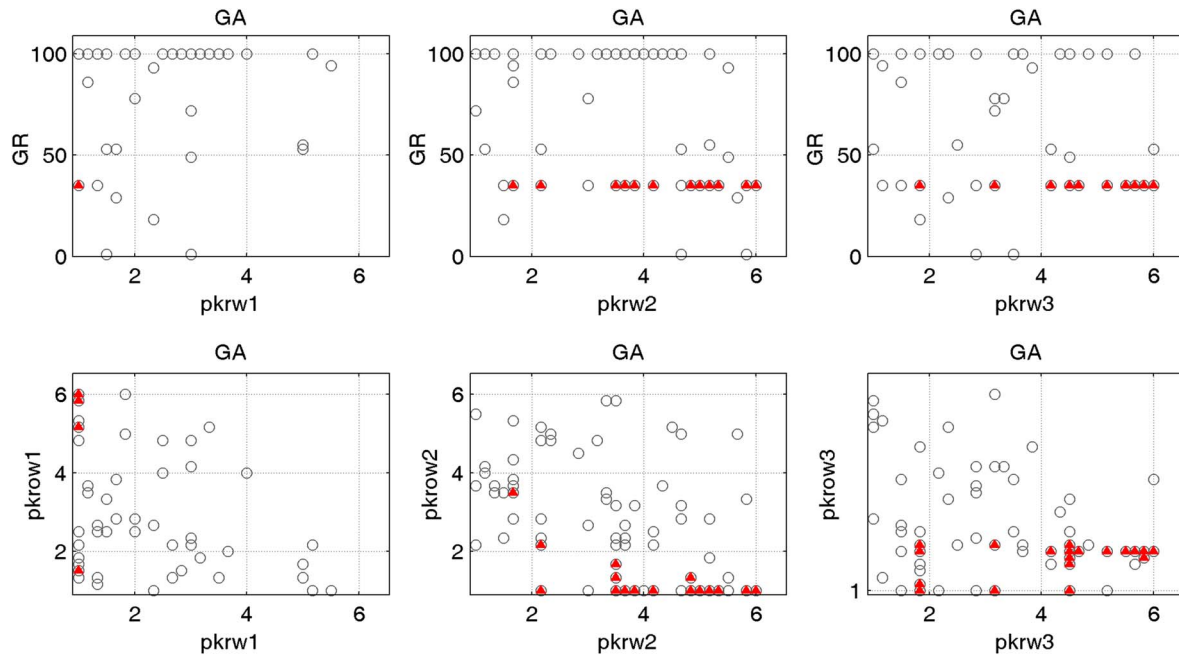


Fig. 16. Cross plots between GR and pk_{rw} (GA). Red points denote models with $|NQDS| \leq 10$ and gray points indicate the models selected according to the procedure described in Section 4.3.

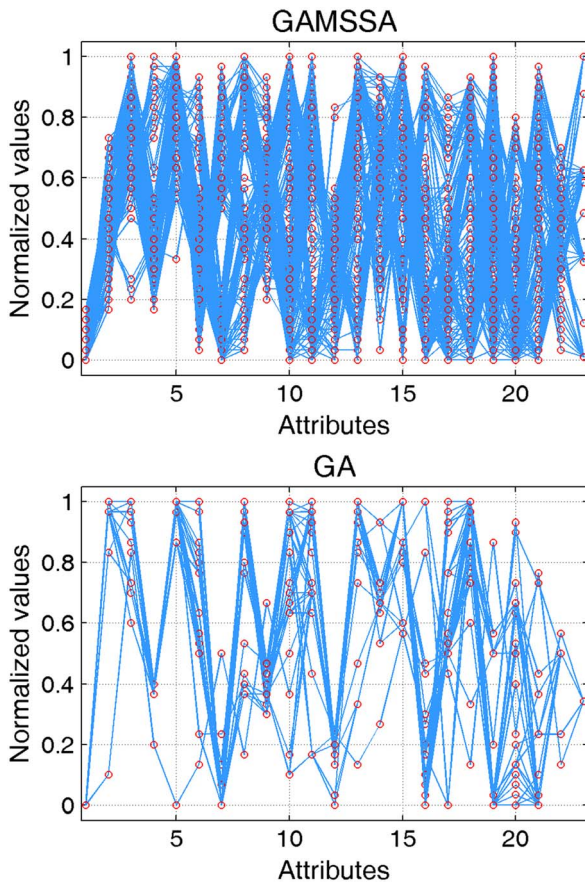


Fig. 17. Parallel coordinate plots for the selected models from GAMSSA and GA.

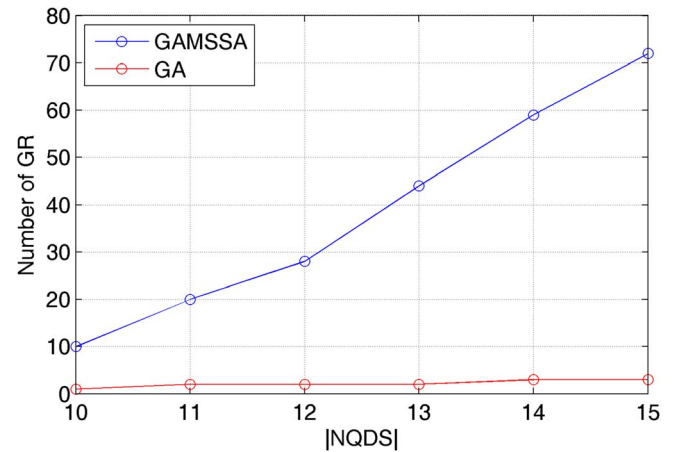


Fig. 18. Variation of the number of geostatistical realization with the NQDS filter.

values are between 4.3 and 5.5. This clearly shows that this is a complex problem characterized by many local minima.

Analyzing the other three plots on the bottom of Figures 15 and 16, we can see a remarkable degree of diversity for the best solutions generated by the proposed method. Note that for facies 1, the best solutions generated by GA collapsed around one single value of pk_{rw1} , while a considerable diversity can be observed for GAMSSA. Larger diversity for GAMSSA is even more pronounced for facies 2 and 3.

Figure 17 shows parallel coordinate plots for the select models from GAMSSA and GA. For visualization purposes, the values of each attribute (represented by red markers)

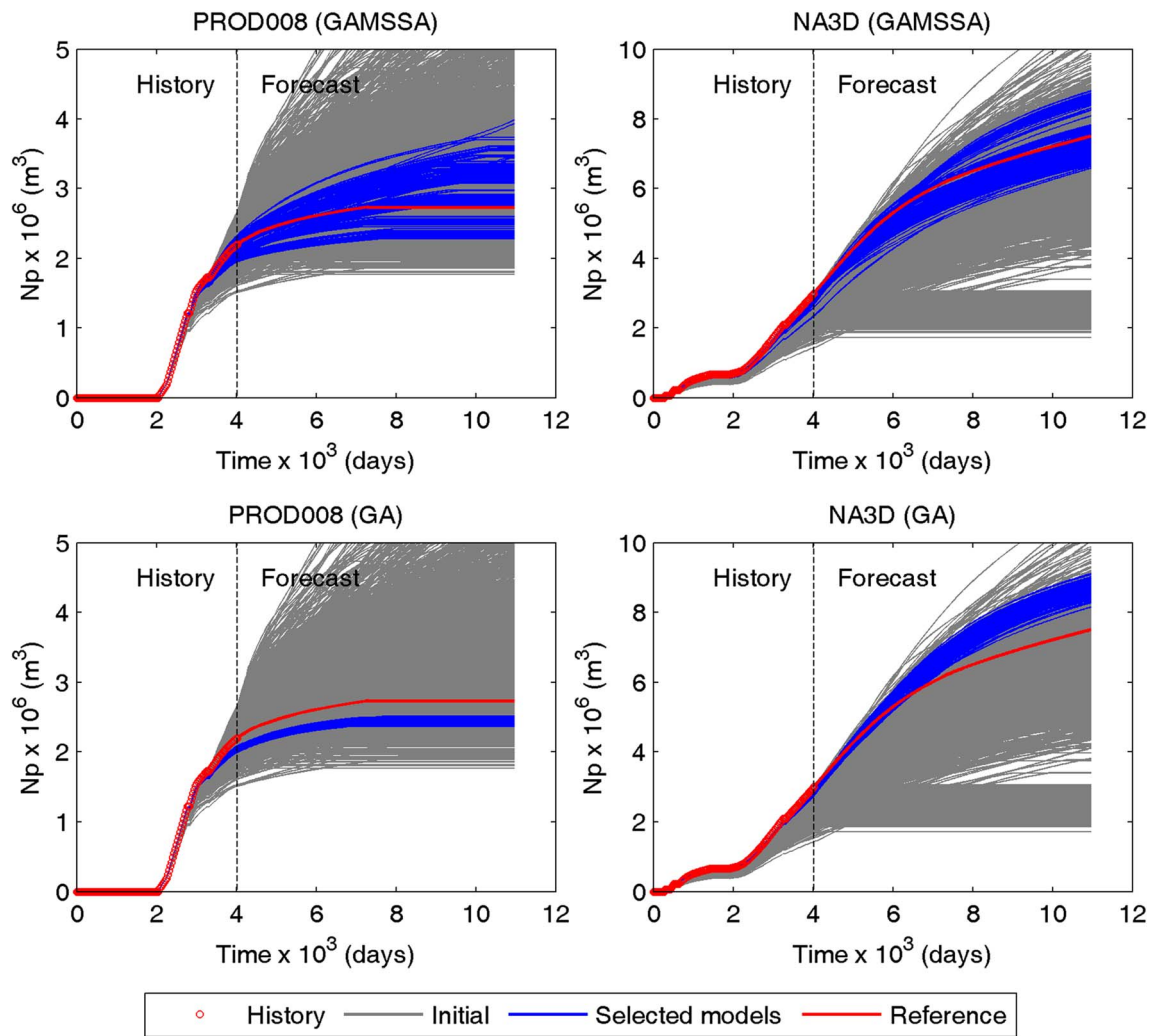


Fig. 19. Cumulative oil production for two producer wells comparing GAMSSA and GA.

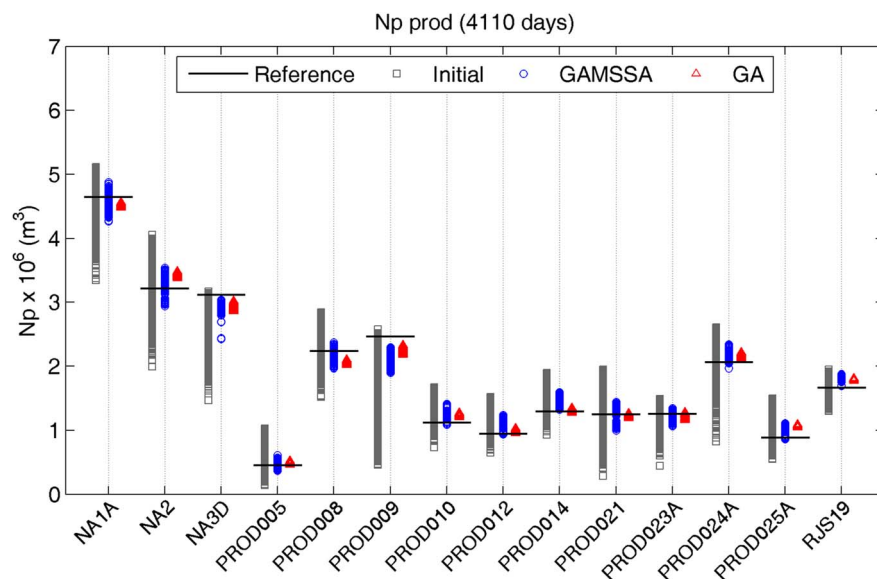


Fig. 20. Distribution of N_p (cumulative oil production) for 4110 days (beginning of forecast) comparing GAMSSA and GA.

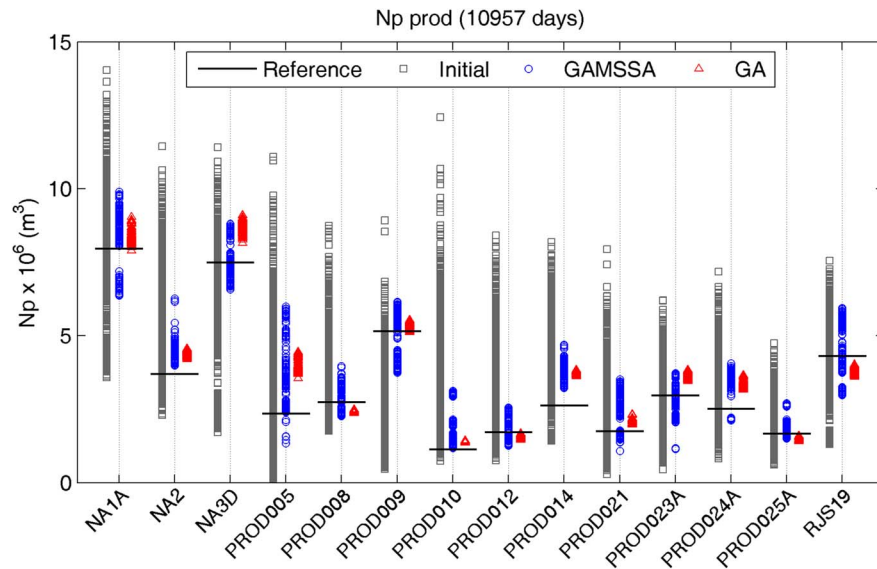


Fig. 21. Distribution of N_p (cumulative oil production) for 10 957 days (end of forecast) GAMSSA and GA.

are mapped into the range between 0 (lower bound) and 1 (upper bound). These plots reveal that the best models selected from the proposed method (GAMSSA) have larger variability compared to the best models selected from GA. Note that for the GA, the models are more concentrated on specific values of the attributes. On the other hand, for GAMSSA the models are more scattered, which denotes larger diversity of the solutions.

Figure 18 shows a comparison between GAMSSA and GA, considering the number of geostatistical realizations versus cut-off value of $|NQDS|$. For a cut-off value of $|NQDS| = 10$, the selected models from GAMSSA contain 10 GR while the selected models from GA contain only one GR. Note also that the number GR increases much faster for GAMSSA than for GA. For $|NQDS| = 15$, for example, there are 72 GR for GAMSSA while for GA there are only three. This analysis reveals that the proposed method preserves the variability of the geostatistical realizations in the solutions, which is very important for more reliable forecast.

4.5 Production forecast

The final goal of the data assimilation process is to improve the predictive capacity of the matched models, increasing the reliability of the decisions made using these models. In this section, we present the results of the production forecast comparing the proposed method and the conventional GA.

Figure 19 shows the cumulative oil production for two producer wells (PROD008 and NA3D). The vertical dashed line divides the history and forecast periods. The gray lines represent the initial variability generated using the 2500 models generated according to the initial uncertainties (see Sect. 3.2). The blue lines represent the selected models and the red line represents the reference model. Clearly, it is possible to note that the solutions provided by GAMSSA encompass the true model, while the solutions from the

GA have lower variability and do not encompass the true solution.

Figures 20 and 21 show, for all producer wells, the distribution of N_p (cumulative oil production) for 4110 days (beginning of forecast) and 10 957 days (end of forecast). First, the variability of N_p from GA is much lower than GAMSSA for all wells. Second, one can observe that the solutions from GAMSSA cover the reference case for most wells, at the beginning and end of the forecast, while the solutions from GA do not encompass the reference case for most wells, especially at the end of the forecast.

5 Final remarks and conclusions

A new optimization scheme using genetic algorithm combined with simulated annealing was developed to integrate geostatistical realizations in data assimilation and reduction of uncertainty process. One of the key advantages of the proposed method is that the geological consistency is maintained. In the case studied, the set of images generated was enough for a good matching. However, if the initial set of realizations is not enough, it is possible to generate more realizations and continue the process.

The proposed scheme takes advantage of the benefits of distributed computing. Genetic algorithm and multi-stat simulated annealing are very suitable for parallel environment. Although the total number of simulations is high, the proposed method is viable because nowadays, computer processors have become cheaper and more widely available. Thus, our methodology has potential for practical applications.

The specific conclusions are:

1. The combination of genetic algorithm with simulated annealing, as proposed in this paper, is an interesting alternative to include geostatistical realizations into

the history matching and reduction of uncertainty process.

2. The variability of the best matched models found using the proposed workflow is higher than the variability of the best models found by the conventional genetic algorithm.
3. The proposed method preserves the variability of the geostatistical realizations in the best solutions. The conventional genetic algorithm tends to lose the diversity of GR, as shown in [Section 4.4](#)
4. The production forecast was strongly affected by the variability of the solutions. The matched models found by the proposed method more properly encompassed the reference case in the forecast period, compared with the matched models found by the conventional genetic algorithm.

Acknowledgments. This work was conducted with the support of Petrobras and *Energi Simulation* within the ANP R&D tax as “commitment to research and development investments.” The authors are grateful for the support of the Center of Petroleum Studies (CEPETRO-UNICAMP/Brazil), the Department of Energy (DE-FEM-UNICAMP/Brazil) and Research Group in *Reservoir Simulation and Management* (UNISIM-UNICAMP/Brazil). In addition, a special thanks to CMG and Schlumberger for the software licenses.

References

- Abdollahzadeh A., Reynolds A., Christie M.A., Corne D. (2012) A parallel GA-EDA hybrid algorithm for history-matching (SPE-153750). *SPE Oil and Gas India Conference and Exhibition*, 28–30 March, Mumbai, India.
- Abdollahzadeh A., Christie M., Corne D., Davies B., Elliott M. (2013) An adaptive evolutionary algorithm for history-matching (SPE-164824). *EAGE Annual Conference & Exhibition Incorporating SPE Europec*, 10–13 June, London, United Kingdom.
- Avansi G.D., Schiozer D.J. (2015) UNISIM-I: Synthetic model for reservoir development and management applications, *Int. J. Model. Simul. Petrol. Ind.* **9**, 1, 21–30.
- Avansi G.D., Maschio C., Schiozer D.J. (2016) Simultaneous history-matching approach by use of reservoir-characterization and reservoir-simulation studies (SPE-179740-PA), *SPE Reserv. Evalu. Eng.* **19**, 4, 69–712.
- Bennett F., Graf T. (2002) Use of geostatistical modeling and automatic history matching to estimate production forecast uncertainty – A case study (SPE 74389). *International Petroleum Conference and Exhibition in Mexico*, 10–12 February, Villahermosa, Mexico.
- Chakra N.C.C., Saraf D.N. (2016) History matching of petroleum reservoirs employing adaptive genetic algorithm, *J. Petrol. Explor. Prod. Technol.* **6**, 4, 653–674.
- Chen P.H., Shahandashti S.M. (2009) Hybrid of genetic algorithm and simulated annealing for multiple project scheduling with multiple resource constraints, *Autom. Constr.* **18**, 434–443.
- Chen C., Jin L., Gao G., Weber D., Vink J.C., Hohl D., Alpak F.O., Pirmez C. (2012) Assisted history matching using three derivative-free optimization algorithms (SPE-154112). *SPE Europec/EAGE Annual Conference*, 4–7 June, Copenhagen, Denmark.
- CMG – Computer Modelling Group Ltd. (2012) *IMEX User’s Guide*, Calgary, Canada.
- Davolio A., Schiozer D.J. (2018) Probabilistic seismic history matching using binary images, *J. Geophys. Eng.* **15**, 261–274.
- Floris F.J.T., Bush M.D., Cuypers M., Roggero F., Syversveen A.R. (2001) Methods for quantifying the uncertainty of production forecasts – a comparative study, *Petrol. Geosci.* **7**, S87–S96.
- Gautier Y., Noetinger B. (2004) Geostatistical parameters estimation using well test data, *Oil Gas Sci. Technol. - Rev. IFP Energies nouvelles* **59**, 2, 167–183.
- Guérillot D., Roggero F. (1995) Matching the future for the evaluation of extreme reservoir development scenarios. *8th European Symposium on Improved Oil Recovery*, 15–17 May, Vienna, Austria.
- Holland J.H. (1975) *Adaptation in natural and artificial system*, The University of Michigan Press, Ann Arbor, MI.
- Hu L.Y., Blanc G., Noetinger B. (2001) Gradual deformation and iterative calibration of sequential stochastic simulations, *Math. Geol.* **33**, 4, 475–489.
- Junghans L., Darde N. (2015) Hybrid single objective genetic algorithm coupled with the simulated annealing optimization method for building optimization, *Energy Build.* **86**, 651–662.
- Kirkpatrick S., Gelatt C.D. Jr, Vecchi M.P. (1983) Optimization by simulated annealing, *Science* **220**, 671–680.
- Li X.G., Wei X. (2008) An improved genetic algorithm-simulated annealing hybrid algorithm for the optimization of multiple reservoirs, *Water Res. Manag.* **22**, 8, 1031–1049.
- Long J., van den Hoek P.J., Alpak F.O., Pirmez C., Fehintola T., Tendo F., Olaniyan E.E. (2012) A comparison of stochastic data-integration algorithms for the joint history matching of production and time-lapse-seismic data (SPE-146418), *SPE Reserv. Evalu. Eng.* **15**, 4, 498–512.
- Maschio C., Vidal A.C., Schiozer D.J. (2008) A framework to integrate history matching and geostatistical modeling using genetic algorithm and direct search methods, *J. Petrol. Sci. Eng.* **63**, 34–42.
- Maschio C., Davolio A., Correia M.G., Schiozer D.J. (2015) A new framework for geostatistics-based history matching using genetic algorithm with adaptive bounds, *J. Pet. Sci. Eng.* **127**, 387–397.
- Maschio C., Schiozer D.J. (2016) Probabilistic history matching using discrete Latin Hypercube sampling and nonparametric density estimation, *J. Petrol. Sci. Eng.* **147**, 98–115.
- Maschio C., Schiozer D.J. (2018) A new methodology for history matching combining iterative discrete Latin Hypercube with multi-start simulated annealing, *J. Petrol. Sci. Eng.* **169**, 560–577.
- Oliveira G.S., Schiozer D.J., Maschio C. (2017) History matching by integrating regional multi-property image perturbation methods with a multivariate sensitivity analysis, *J. Petrol. Sci. Eng.* **153**, 111–122.
- Ouenes A., Bhagavan S., Bunge P.H., Travis B.J. (1994) Application of simulated annealing and other global optimization methods to reservoir description: Myths and realities (SPE-28415). *SPE Annual Technical Conference and Exhibition*, 25–28 September, New Orleans, Louisiana.
- Roggero F., Guérillot D. (1996) Gradient method and Bayesian formalism application to petrophysical parameter

- characterization. *5th European Conference on the Mathematics of Oil Recovery*, 3–6 September, Leoben, Austria.
- Romero C.E., Carter J.N., Zimmerman R.W., Gringarten A.C. (2000) Improved reservoir characterization through evolutionary computation (SPE-62942). *Annual Technical Conference and Exhibition*, 1–4 October, Dallas, Texas.
- Romero C.E., Carter J.N. (2001) Using genetic algorithms for reservoir characterization, *J. Petrol. Sci. Eng.* **31**, 2–4, 113–123.
- Sanghyun L., Stephen K.D. (2018) Optimizing automatic history matching for field application using genetic algorithm and particle swarm optimization (OTC-28401-MS). *Offshore Technology Conference Asia*, 20–23 March, Kuala Lumpur, Malaysia.
- Sayyafzadeh M., Haghighi M., Carter J.N. (2012) Regularization in history matching using multi-objective genetic algorithm and Bayesian framework (SPE-154544). *SPE Europec/EAGE Annual Conference*, 4–7 June, Copenhagen, Denmark.
- Schulze-Riegert R.W., Axmann J.K., Haase O., Rian D.T., You Y.L. (2002) Evolutionary algorithms applied to history matching of complex reservoirs (SPE-77301), *SPE Reserv. Evalu. Eng.* **5**, 2, 163–173.
- Schulze-Riegert R.W., Haase O. (2003) Combined global and local optimization techniques applied to history matching (SPE-79668). *Reservoir Simulation Symposium*, 3–5 February, Houston, Texas.
- Sultan A.J., Ouenes A., Weiss W.W. (1994) Automatic history matching for an integrated reservoir description and improving oil recovery (SPE-27712). *Permian Basin Oil and Gas Recovery Conference*, 16–18 March, Midland, Texas.
- UNISIM (2015) *Research group on Reservoir Simulation and Management*. UNISIM-I-H: Case Study for History Matching. <https://www.unisim.cepetro.unicamp.br/benchmarks/en/unisim-i/unisim-i-h>.
- Xavier C.R., dos Santos E.P., da Fonseca Vieira V., dos Santos R.W. (2013) Genetic algorithm for the history matching problem, *Procedia Comput. Sci.* **18**, 946–955.